

What Is Claimed Is:

1. A method for filtering an electronic document to determine whether content in the electronic document is inappropriate to users, the method comprising:

    parsing the electronic document to extract the relevant words in the document;

    assigning a weight to each relevant word in the document;

    extracting a plurality of contexts for each relevant word in the document from a thesaurus dictionary;

    assigning a weight to each context in the plurality of contexts;

    determining which contexts in the plurality of contexts are the most important contexts in the document; and

    restricting access to the electronic document if the most important contexts in the document are in a list of restricted contexts.

2. The method of claim 1, further comprising restricting access to the electronic document if the electronic document is a web page and the web page is in a list of restricted web pages.

3. The method of claim 1, wherein assigning a weight to each relevant word in the document comprises assigning a weight according to one or more formatting parameters selected from a group of formatting parameters consisting of: number of times the relevant word appears in the document; total number of words in the document; format of the relevant word in the document; format of a

plurality of words surrounding the relevant word in the document; header or meta tag associated with the relevant word if the electronic document is a web page; and PICS rating associated with the document.

4. The method of claim 1, wherein extracting a plurality of contexts for each relevant word in the document from a thesaurus dictionary comprises creating a context vector for each relevant word in the document comprising the plurality of contexts found in the thesaurus dictionary.

5. The method of claim 1, wherein assigning a weight to each context in the plurality of contexts comprises determining the number of words in the document having the same context and the number of contexts associated with each word in the document.

6. The method of claim 5, wherein the weight is based on the weight of the relevant word; the number of words in the document having the same context; and the number of contexts associated with each word in the document.

7. The method of claim 1, wherein determining which contexts in the plurality of contexts are the most important contexts in the document comprises determining which contexts in the plurality of contexts have the highest weight.

8. The method of claim 1, wherein restricting

access to the electronic document if the most important contexts in the document are in a list of restricted contexts comprises displaying a message to the user notifying the user that the document has inappropriate content.

9. A method for filtering an electronic document to determine whether content in the electronic document is inappropriate to users, the method comprising:

    checking whether the electronic document is in a list of restricted electronic documents;

    determining whether the electronic documents contains an unacceptable number of inappropriate words or pictures;

    extracting a plurality of contexts for each word in the document from a thesaurus dictionary;

    assigning a weight to each context in the plurality of contexts;

    determining which contexts in the plurality of contexts are the most important contexts in the document; and

    restricting access to the electronic document if the most important contexts in the document are in a list of restricted contexts.

10. The method of claim 9, wherein the electronic document comprises one or more electronic documents selected from a group consisting of: a web page; a newsgroup transcript; a chat room transcript; an e-mail; a document in a CD; a document in a DVD; and a document in a disk.

11. The method of claim 9, wherein determining whether the electronic documents contains an unacceptable number of inappropriate words or pictures comprises determining a ratio of pictures to words in the document and determining the number of inappropriate words in a plurality of links in the document if the ratio exceeds fifty percent.

12. The method of claim 9, wherein assigning a weight to each context in the plurality of contexts comprises determining the number of words in the document having the same context and the number of contexts associated with each word in the document.

13. The method of claim 9, wherein determining which contexts in the plurality of contexts are the most important contexts in the document comprises determining which contexts in the plurality of contexts have the highest weight.

14. A system for filtering an electronic document to determine whether content in the electronic document is inappropriate to users, the system comprising:

a configuration user interface for allowing a filtering software administrator to control the users' access to electronic documents;

a filtering software plug-in to monitor users' access to electronic documents;

an Internet sites database storing a list of inappropriate sites;

a context database storing a list of restricted contexts; and

a thesaurus database storing a thesaurus dictionary.

15. The system of claim 14, wherein the the electronic document comprises one or more electronic documents selected from a group consisting of: a web page; a newsgroup transcript; a chat room transcript; an e-mail; a document in a CD; a document in a DVD; and a document in a disk.

16. The system of claim 14, wherein the configuration user interface comprises a user interface for specifying which sites and contexts are inappropriate to users.

17. The system of claim 14, wherein the filtering software plug-in performs a contextual analysis of the electronic document to determine whether the electronic document is inappropriate to users.

18. The system of claim 17, wherein the contextual analysis comprises determining the main contexts of the electronic document.

19. The system of claim 18, wherein the main contexts of the electronic document comprise the contexts assigned a higher weight.

20. The system of claim 19, wherein the weight comprises a value assigned to a context extracted from the thesaurus database, the value depending on one or more parameters selected from a group of parameters consisting

of: number of words having the same context; weights of the words having the same context; and number of words in the document.